

DOCUMENT RESUME

ED 392 839

TM 024 617

AUTHOR Kim, JinGyu; McLean, James E.
TITLE The Influence of Examinee Test-Taking Motivation in Computerized Adaptive Testing.
PUB DATE Apr 95
NOTE 22p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (San Francisco, CA, April 19-21, 1995).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Ability; *Adaptive Testing; Aptitude Tests; *College Students; *Computer Assisted Testing; Foreign Countries; Higher Education; Mathematics Achievement; Multivariate Analysis; Self Concept; Student Attitudes; *Student Motivation; *Test Anxiety
IDENTIFIERS Korea; *Koreans; Paper and Pencil Tests

ABSTRACT

The purpose of the study was to investigate the effects of test motivation on estimated ability, test anxiety, and attitudes toward computerized adaptive testing (CAT). Korean college students (n=208) were given the Math Aptitude Test, Math Self-Concept Scale, Math Test Anxiety Scale, Computer Competence Instrument, Computer Anxiety Scale, and Test Anxiety Inventory in the regular classroom. The two groups (motivated and non-motivated) were randomly assigned by each course section. The motivated group was given special test instructions. The paper-and-pencil test (PPT) and the CAT algebra tests were given to each group in random order (PPT-CAT or CAT-PPT) under the counterbalanced design at the computer laboratory. They were also given a 10-item paper test anxiety scale, a 10-item computer test anxiety scale, and a paper-and-pencil version of the Questionnaire on Computerized Adaptive Testing. A multivariate analysis of covariance, with the math aptitude and the test anxiety as covariates, demonstrated that test motivation influenced improvement in estimated ability and reduction in test anxiety, but did not affect CAT attitudes. (Contains 4 tables and 45 references.) (Author)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

JINGYU KIM

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

The Influence of Examinee Test-Taking Motivation in Computerized Adaptive Testing

JinGyu Kim and James E. McLean
The University of Alabama

◆◆◆

Paper Presented at the Annual Meeting of the
National Council on Measurement in Education
San Francisco, CA
April 1995

BEST COPY AVAILABLE

Abstract

The purpose of the study was to investigate the effects of test motivation on estimated ability, test anxiety, and attitudes toward computerized adaptive testing (CAT). Korean college students ($n=208$) were given the Math Aptitude Test, Math Self-Concept Scale, Math Test Anxiety Scale, Computer Competence Instrument, Computer Anxiety Scale, and Test Anxiety Inventory in the regular classroom. The two groups (Motivated and Non-motivated) were randomly assigned by each course section. The motivated group was given special test instructions. The paper-and-pencil test (PPT) and the CAT algebra tests were given to each group in random order (PPT-CAT or CAT-PPT) under the counterbalanced design at the computer laboratory. They were also given a 10-item paper test anxiety scale, a 10-item computer test anxiety scale, and a paper-and-pencil version of the Questionnaire on Computerized Adaptive Testing. A multivariate analysis of covariance, with the math aptitude and the test anxiety as covariates, demonstrated that test motivation influenced improvement in estimated ability and reduction in test anxiety, but did not affect CAT attitudes.

The Influence of Examinee Test-Taking Motivation in Computerized Adaptive Testing*

Computerized adaptive testing (CAT) has several advantages over conventional paper-and-pencil testing (PPT). Most CAT studies have tried to demonstrate the efficiency and precision of the examinee's ability estimation in CAT over PPT (Bergstrom & Lunz, 1992; Betz & Weiss, 1973, 1974; Frick, 1989, 1990; Kingsbury, 1990; McBride & Martin, 1983; McKinley & Reckase, 1980; Olsen, Maynes, Slawson, & Ho, 1989; Weiss, 1982). However, research on the effects of demographic and psychological characteristics of examinees on CAT has been largely neglected (Legg & Buhr, 1992; Kim, 1993). In other words, how examinees' individual differences are systematically related to CAT methods has not been studied extensively.

Recently, a few researchers have paid attention to individual differences in CATs (Legg & Buhr, 1992; Rocklin & O'Donnell, 1987; Vispoel & Rocklin, 1993). Their studies investigated the effects of several individual difference variables (age, ability, computer anxiety, computer experience, gender, math self-concept, and test anxiety) on test performance. They also explored whether group differences in reactions to computerized test administration could help explain observed differences between CAT and PPT (Legg & Buhr, 1992), between computerized test (CT) and self-adapted test (Rocklin & O'Donnell, 1987), and among CT, CAT, and self-adapted testing (Vispoel & Rocklin, 1993). Previous studies demonstrated the efficiency and precision of CAT, but showed inconclusive results with

* The authors wish to thank Drs. Steven L. Wise and Barbara S. Plake who allowed us to use a set of algebra skills items they developed for use in another study.

respect to specific individual difference variables and the relationships between individual difference variables and test performance.

Few studies have compared individual differences in student test-taking motivation between PPT and CAT. Generally, researchers use many kinds of rewards to motivate students to pay attention to their tasks, to use their full abilities, and to enlist their best efforts in doing their work. Student test-taking motivation is a possible factor in the performance of students on nearly all standardized achievement tests (Brown & Walberg, 1993), large-scale state-constructed achievement tests (Linn, 1993; Wolf, 1993; Wolf, Smith, & Birnbaum, 1994), and international achievement tests (Wainer, 1993). The content of those tests is often unrelated to specific topics that students have recently studied, and their performance on such tests does not affect their grades. Wainer (1993) has argued that the outstanding performance of Korean students in two international comparative studies may be due to their high student test-taking motivation because the chosen students are representing the honor of their school and their country in this competition.

Brown and Walberg (1993) found the motivational effects on test scores of elementary students, using special test instruction enhanced their test-taking motivation. Wolf (1993) found that motivation to perform on a test is directly influenced by the consequence of the level of performance, and that motivation to perform enhances test performance. Wolf, et al. (1994) proposed an extension of Pintrich's expectancy theory of motivation (Pintrich, 1988, 1989) in which the expectancy component includes not only the individual's assessment of how likely he or she is to do well on an item, but also how much work will be involved in the completion of the item. They also found that an item that required much time required more motivation to put forth the mental energy necessary to arrive at a correct answer. Kim (1995) found that examinees' test motivation is a significant predictor relating to computerized

adaptive test performance. Therefore, we can say that examinees' test motivation is associated with higher level of test performance on both PPTs and CATs.

Previous research on the relationship between test anxiety and achievement motivation offered a variety of interpretation (Atkinson, 1964; Hill, 1972, 1980; Sarason, Davidson, Lighthall, Waite, & Ruebush, 1960). Hill (1972) combined elements of both the Sarason theory of anxiety and the Atkinson approach to achievement motivation. Hill's conceptualization agrees with Sarason's theory in that evaluative reactions from adults are believed to underlie and enhance the effects of success/failure experiences. It also utilizes Atkinson's theory of emphasizing motives to approach success and avoid failure and trying to delineate the absolute strength of the two motives (Phillips, Pitcher, Worshman, & Miller, 1980). McMillan (1980) explained that success is usually associated with a positive affect that alleviates anxiety, whereas failure tends to elicit a negative feeling that increases anxiety. Mevarech and Ben-Artzi (1987) demonstrated that computer-assisted instruction (CAI) as compared to traditional instruction tends to yield lower scores on various aspects of mathematics anxiety. According to Mevarech and Ben-Artzi (1987), the high rate of success/failure provided in CAI settings contributed to lower levels of worries about learning mathematics. Like CAI settings, we may expect a lower level of test anxiety in CAT settings than in PPT conditions because the rate of success/failure is usually not controlled. If examinees' test motivation is associated with higher levels of test performance in CAT test situations, examinees may have a higher rate of success and sufficient motivation to alleviate anxiety. It is worthwhile to investigate the effects of test motivation not only on test performance but also on test anxiety and attitudes in CAT test conditions.

The purpose of this study was to examine the effects of examinees' test-taking motivation on estimated ability, test anxiety, and attitude toward computerized adaptive testing to identify the feasibility of the CAT in college classroom testing.

Method

Subjects and Treatments

The subjects in this study were a sample of 208 Korean college students enrolled in Test & Measurements courses at the Kyungpook National University, Taegu, Korea. Data were collected from six course sections during the spring semester of 1994.

Two treatments, the test method and student test-taking motivation level, were used as independent variables in this study. The paper-and-pencil test (PPT) method was a designed test of 20 items selected from a pool of 70 algebra items (Wise, Plake, Johnson, & Roos, 1992). The computerized adaptive test (CAT) method used a pool of 70 algebra items. The adaptive test terminated when 20 items were administered or when the variance of .10 was reached (Legg & Buhr, 1992), whichever come first.

The motivated subject group was given special test instructions that indicated the test was very important because it was helpful to learn concepts of basic statistics, they would receive extra points, and results would be used to compare Korean college students with their American counterparts. The non-motivated group was given only the standard test instructions.

Instruments

Algebra Item Bank. A pool of items was developed to be used in identifying those introductory statistics students whose basic mathematics skills were in need of remediation (Wise, et al., 1992). After reviewing the 90 algebra test items, the authors of the current study found that 70 items could apply to Korean students without translation. The feedback

statements that are usually a part of CAT were deleted to make the test conditions as equal as possible with those for the PPT. The CAT algebra test was administered on IBM microcomputers using the MicroCAT Testing System (Assessment Systems Corporation, 1989).

Math Aptitude Test. A math aptitude test was used in this study. The test consisted of 20 retired Graduate Record Examination items. The two Korean mathematics professors and the senior author chose five items in each of four subscales. Each item used a five-option, multiple-choice format. The test was translated into Korean by the first author. Once the Korean version had been prepared, the two Korean mathematics professors were asked to check the test items and compare the Korean version to the English version.

Computer Competence Instrument. The Computer Competence Instrument, developed by Martinez and Mead (1988), was used to measure test takers' computer vocabulary and their knowledge about computers. The instrument consisted of 33 four-choice items.

Computer Anxiety Scale. The computer anxiety subscale from the Computer Attitude Scale (Loyd & Gressard, 1984) was used to measure computer anxiety. The subscale had 10 four-option agree-disagree items. The range of the alpha reliabilities in the 10 previous studies was from .57 to .93, and the median was .87 (Kim, McLean, & Moon, 1994).

Test Anxiety Inventory. The original Test Anxiety Inventory (TAI; Spielberger, 1980) was used to measure general test anxiety one week before examinees took the algebra skills test. The modified TAI was used to measure paper and computer test anxiety after taking PPT and after taking CAT. The original TAI scale consisted of 20 four-option Likert-type items (1=almost never, 4=almost always). The general test anxiety scale used the original 20 items, but the paper test anxiety scale changed the word "test" into "paper-and-pencil test."

The computer test anxiety scale changed the word "test" into "computer adaptive test." The paper and computer test anxiety scales chose and used 10 of the original 20 items that specifically asked how test-takers felt during test administration. The 10 items were used to measure computer test anxiety (Powell, 1991). The range of possible scores for the modified TAI scale varied from a minimum of 10 to a maximum of 40.

Questionnaire on Computerized Adaptive Testing. A questionnaire was used to assess attitudes toward computerized adaptive tests. The questionnaire was designed to identify the examinees' opinions about computerized adaptive testing. The scale consisted of 20 items and was a four-point Likert-type measure. Ten items were chosen from the Questionnaire on Computer-Assisted Testing (Legg & Buhr, 1992) and five items were modified from the Evaluation Questionnaire (Schmidt, Urry, & Gugel, 1978). Five items were added by the authors. A principal component analysis of the Attitudes toward Computerized Adaptive Testing Questionnaire was conducted, followed by a factor analysis using a two-factor solution with a varimax rotation. The first factor item loadings were related to general characteristics of computerized testing and suggested G-factor CAT attitudes. The second factor item loadings were related to specific characteristics of computerized adaptive testing to compare those of paper-and-pencil testing, and suggested S-factor CAT attitudes.

The translation of the instruments into Korean. All of the instruments, except the Math Aptitude Test, were translated into Korean using the same procedures. The first author translated into Korean; every attempt was made to provide a Korean version that was as faithful a representation of the English as possible. The Korean versions of the instruments were translated back into English by a Korean linguist, and the back-translated versions were compared with the original English to ensure that the translation was accurate. A Korean

educational psychologist and the author checked the Korean and the back-translated English versions of the instruments and compared them to the original English versions.

Procedures

The students took the four instruments (Math Aptitude Test, Computer Competence Instrument, Test Anxiety Inventory, and Computer Anxiety Scale) under the supervision of their instructors in the first week of class. Instruments were assembled into packets and distributed in a random fashion to the students as they arrived for class. Each packet contained the four measures arranged in random order. The course instructors notified their students that at a designated time in the following week they would be taking a computerized test in the computer laboratory instead of hearing a lecture.

During the second and third weeks of the semester, two special computer rooms which shared a controllable box were designated for use by the course sections of this study in taking a computerized test. The two groups (Motivated and Non-motivated) were randomly assigned by each course section. Three course sections were the motivated subject group, and the other three course sections were the non-motivated subject group. To control the effect of the test motivation, the non-motivated subject groups took the test first, followed by the motivated subject groups.

As students arrived at the computer laboratory for testing, they were randomly assigned to one of two groups (PPT first and CAT first), each meeting in a different room. The students were seated at one of 44 IBM 486 model microcomputers. Two algebra tests were given to each subject group in random order (PPT-CAT or CAT-PPT) in each room. The test administration procedures of each group were as follows: The group taking the PPT first were given the PPT test and the 10-item paper test anxiety scale, and then were given the CAT sample test of three example items. After completing the sample test, they took the main

CAT and the 10-item computer test anxiety scale. The group taking CAT first were given the same tests in the opposite order. Finally, on completion of the two algebra tests and the two test anxiety scales, students were given a paper-and-pencil version of the Questionnaire on Computerized Adaptive Testing.

Data Analysis

The examinees' test-taking motivation (motivated, non-motivated) was the independent variable in a multivariate analysis of covariance (MANCOVA). There were six dependent variables of interest in this study: (a) two estimated algebra skills (PPT and CAT), (b) two test anxiety scores (PPT test anxiety and CAT test anxiety), and (c) two CAT attitudes scores (G-factor and S-factor). The covariates were the math aptitude and the test anxiety.

Results

Preliminary Analyses

The linear relationships between the covariates (math aptitude and test anxiety) and the dependent variables (PPT algebra skills, CAT algebra skills, paper test anxiety, and computer test anxiety) were conducted by the multivariate test for significant regression. The results showed that there were significant relationships--Wilks' Lambda = .5130, $F(2, 413) = 196.02$, $p = .0001$ for math aptitude and Wilks' Lambda = .8946, $F(2, 413) = 24.33$, $p = .0001$ for test anxiety. Thus, the results indicated that the covariates were related to dependent variables.

The tests of the homogeneity of regression for the present data failed to yield significant F ratios--Wilks' Lambda = .9899, $F(2, 203) = 1.03$, $p = .3577$ for math aptitude by test motivation interaction and Wilks' Lambda = .9843, $F(2, 203) = 1.62$, $p = .2005$ for test anxiety by test motivation interaction. Thus, there was no significant interaction between the covariate (math aptitude) and the independent variable (test motivation); there was also no

significant interaction between the covariate (test anxiety) and the independent variable (test motivation). Therefore, the regression lines for each group were assumed to be parallel.

To identify the base line for examinees' individual differences, which they had before the treatments were given, a series of t -tests were conducted by test motivation group. The results showed that the mean scores for the motivated group examinees and the non-motivated group examinees did not differ significantly in the major individual difference variables (math aptitude, computer literacy, and test anxiety; $p > .05$) except for the computer anxiety variable ($p < .05$).

Descriptive Statistics

Means, standard deviations, reliability coefficients, and correlations for the main variables are presented in Table 1.

Table 1
Descriptive Statistics for Major Variables

Variables	M	SD	r	Variables					
				1	2	3	4	5	6
1. Math Aptitude	14.59	3.01	.74		.45*	-.05	-.12	.59*	.63*
2. Computer Literacy	23.41	4.98	.80			-.26*	-.06	.25*	.37*
3. Computer Anxiety	23.68	5.66	.87				.27*	-.01	-.05
4. Test Anxiety	38.46	8.35	.90					-.06	-.13
5. PPT Algebra Skills	1.81	.96	.62						.54*
6. CAT Algebra Skills	1.79	.77	.74						

Note. r , reliabilities are alpha coefficients. PPT algebra skills were determined by a special data set and the conventional test scoring program (SCOREALL) of the MicroCAT testing system. CAT algebra skills were determined by the examination subsystem of the MicroCAT testing system. Reliability estimates for the PPT and the CAT were derived by subtracting the mean variance error of estimated ability from one (Urry, 1977) using both a fixed item (20 items) and a .10 variance error termination rule. Variance errors for PPT were computed by a special computer program (Yoes, 1994).

* The correlation coefficients over .14 are significant at the .05 level.

The mean and standard deviation for PPT algebra skills test were similar to those for CAT algebra skills test. Reliability estimates for the measures ranged from .62 for the PPT algebra skills test to .90 for the Test Anxiety Inventory. The algebra skill-related measures (math aptitude and algebra skills tests) were positively correlated, computer literacy and algebra skill-related measures were positively correlated, test anxiety and computer anxiety were positively correlated, and computer literacy and computer anxiety were negatively correlated.

Group Differences

Means and standard deviations for test motivation on estimated algebra skills, test anxiety, and CAT attitudes are presented in Table 2.

Table 2

Means and Standard Deviations for Estimated Algebra Skills, Test Anxiety, and CAT Attitudes by Test Motivation Groups

Variables	Motivated (<i>n</i> =104)		Non-Motivated (<i>n</i> =104)		Total (<i>n</i> =208)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Algebra Skills						
PPT	1.94	0.96	1.68	0.95	1.81	0.96
CAT	1.92	0.83	1.67	0.68	1.79	0.77
Test Anxiety						
PPT	11.36	2.32	13.01	3.98	12.37	3.31
CAT	13.76	4.28	14.13	4.32	13.74	4.31
CAT Attitudes						
G-Factor	33.32	3.96	32.27	4.03	32.80	4.02
S-Factor	26.04	5.30	24.44	4.90	25.24	5.15

Group differences in estimated algebra skills. Group differences in estimated algebra skills were analyzed with a multivariate approach using a mixed-factor model. The one between-subjects factor was test motivation (motivated group and non-motivated group) and

the within-subjects factor was test method (PPT and CAT). The 2 x 2 multivariate repeated measures analysis of covariance, in which the covariate was math aptitude, revealed only one significant main effect for test motivation group (Wilks' Lambda = .9460, $F(1, 205) = 11.69$, $p = .0008$). Neither the test method main effect (Wilks' Lambda = .9911, $F(1, 205) = 1.83$, $p = .1772$) nor the test motivation group x test method interaction (Wilks' Lambda = .9999, $F(1, 205) = .01$, $p = .9390$) was significant.

Table 3 provides the univariate ANCOVA summary with repeated measures in estimated algebra skills.

Table 3

Univariate ANCOVA Summary for Estimated Algebra Skills

Source	df	SS	MS	F	p
Covariate					
Math Aptitude ^a	1	114.89	114.89	246.82	.0001
Between Subjects					
Motivation (M)	1	6.72	6.72	11.69	.0008
Subject (M)	205	117.93	.58		
Within Subjects					
Test Method (T)	1	.65	.65	1.83	.1772
M x T	1	.01	.01	.01	.9390
T x Subject(M)	206	73.40	.36		

Note: ^a Based on pooled error term $SS_{Cov\ Error} = SS_{s(M)} + SS_{T \times s(M)} / df_{s(M)} + df_{T \times s(M)} = .47$.

As can be seen in Table 3, the results indicated significant differences for the test motivation group ($F(1, 205) = 11.69$, $p = .008$) and covariate (math aptitude). The test method and test method x test motivation interaction indicated no significant differences.

Motivated group examinees had higher estimated ability than did non-motivated group examinees. There was no significant difference between PPT and CAT in estimated ability.

Group differences in test anxiety. Group differences in test anxiety were analyzed with a multivariate analysis of a mixed-factor model. The one between-subjects factor was test motivation (motivated group and non-motivated group) and the within-subjects factor was test method (PPT and CAT). The 2 x 2 multivariate repeated measures analysis of covariance, in which the covariate was general test anxiety, revealed only one significant main effect for the test motivation group (Wilks' Lambda = .9714, $F(1, 205) = 6.02$, $p = .0149$). There were no significant differences in test method main effect (Wilks' Lambda = .9973, $F(1, 205) = .55$, $p = .4600$) or in the test motivation group x anxiety type interaction (Wilks' Lambda = .9962, $F(1, 205) = .79$, $p = .3759$).

Table 4 provides the univariate ANCOVA summary with repeated measures in test anxiety.

Table 4
Univariate ANCOVA Summary for Test Anxiety

Source	df	SS	MS	F	p
Covariate					
Test Anxiety ^a	1	438.74	438.74	32.44	.0001
Between Subjects					
Motivation (M)	1	111.85	111.85	6.02	.0149
Subject (M)	205	3806.43	18.57		
Within Subjects					
Test Method (T)	1	4.68	4.68	.55	.4600
M x T	1	6.75	6.75	.79	.3759
T x Subject(M)	206	1751.50	8.50		

Note: ^a Based on pooled error term $SS_{Cov\ Error} = SS_{s(M)} + SS_{T \times s(M)} / df_{s(M)} + df_{T \times s(M)} = 13.52$.

As can be seen in Table 4, the results indicated significant differences for the test motivation group ($F(1, 205) = 6.02, p = .0149$) and the covariate (general test anxiety). There were no significant differences in test method and in the test method x test motivation interaction. Non-motivated group examinees had higher test anxiety scores than did motivated group counterparts. There were no significant differences between the CAT and the PPT test method based on test anxiety.

Group differences in attitudes toward computerized adaptive testing. Group differences in attitudes toward computerized adaptive testing were analyzed using a one-way multivariate analysis of variance (MANOVA). The test motivation was an independent variable and two factors of the attitudes toward CAT were dependent variables. The MANOVA indicated non-significant main effects for the test motivation group (Wilks' Lambda = .9723, $F(2, 202) = 2.88, p = .058$).

Motivated group examinees had more positive attitudes toward CAT than did their non-motivated counterparts, but did not result in a statistically significant difference. The results indicated that test motivation group examinees had similar attitudes toward CAT.

Discussion

The purpose of the study was to investigate the effects of test motivation on estimated ability, test anxiety, and attitudes toward computerized adaptive testing for Korean college students. Motivated examinees had higher mean ability estimates than non-motivated examinees in this study. The results found that test motivation affected the facilitation of algebra skills estimates significantly. The present results are consistent with findings from previous studies that investigated motivational effects on paper-and-pencil test scores (Brown & Walberg, 1993; Wolf, et al., 1994). Non-motivated examinees had higher test anxiety

levels than did their motivated counterparts. Although motivated examinees had higher computer anxiety levels than non-motivated examinees before the treatments were given, the results showed that test motivation affected the reduction of examinees' test anxiety.

There is at least one possible explanation for this result. Motivated examinees had higher scores than did non-motivated examinees in the current study and in previous studies (Brown & Walberg, 1993; Wolf, 1993). Motivated examinees might have experienced a higher level of initial success on the tests than did non-motivated counterparts. It was found that test anxiety increased in students who experienced initial failure in a testing context, while test anxiety decreased in students who experienced initial success (Bradshaw & Gaudry, 1968; McMillan, 1980; Mevarech & Ben-Artzi, 1987). Perhaps this initial success may make motivated examinees feel comfortable and more responsive to positive statements.

There was no significant difference between motivated examinees and non-motivated examinees in attitudes toward CAT. The present finding is not consistent with previous CAT research in which examinees generally preferred computerized testing formats (Gershon & Bergstrom, 1991; Legg & Buhr, 1992; Moe & Johnson, 1988; Schmitt, et al., 1978; Vispoel & Coffman, 1992).

The current data show that test motivation was related to an improvement in examinees' estimated ability and a reduction in test anxiety. These findings provide empirical evidence that the effects of test motivation on test performance and test anxiety appear in CAT as well as in PPT.

The present results have useful practical implications for educators and researchers. First, test motivation is one of the important factors which explains U.S. students' poor performance on achievement relative to students in other countries (Brown & Walberg, 1993). The Korean students' exceptional math performance in international comparison

studies is attributable to their higher motivation (Wainer, 1993) or more time spent doing math homework (Lapointe, Mead, & Phillips, 1989). Moon, Byun, McLean, and Kaufman (1994) found that Korean children's math strength may have influenced their exceptional ability to remember numbers rather than vice versa. Also, they suggest that higher motivation may have influenced their high math achievement even more so than their outstanding sequential processing.

Second, the present data provide an explanation of the relationship between test motivation and test anxiety. These motivational effects on test anxiety might have a practical implication for educators. Highly motivated test instructions might contribute to reducing examinees' test anxiety rather than to increasing it. Thus, educators must know how to use motivated test instructions for their classroom tests.

Finally, this study is one of the first attempts at using non-English speaking subjects in CAT research and provides an empirical foundation for further cross-cultural studies in CAT.

The study was limited in the following areas. This study did not measure the degree of examinees' initial success or their satisfaction with test results after taking a test. To explain the relationship between test motivation and test anxiety systematically, further research should be considered concerning the measurement of the examinees' initial success on test scores. In this study, the levels of examinees' test motivation were not measured before or after test instructions were given. A procedure to measure examinee test motivation should be developed to assess the more genuine effects of test motivation on test performance, test anxiety, and attitudes. Further international comparison studies are needed that control the effects of test motivation on test performance, using examinees' test motivation measures. Moreover, further CAT research should be focused on how to enhance students' motivation not only in tests, but also in classroom learning and instruction.

References

- Assessment Systems Corporation (1989). User's manual for the MicroCAT testing system. version 3. St. Paul, MN: Author.
- Atkinson, A.W. (1964). An introduction to motivation. Princeton, NJ: Van Nostrand.
- Bergstrom, B.A., & Lunz, M.E. (1992). Confidence in pass/fail decisions for computer adaptive and paper and pencil examinations. Evaluation and The Health Professions, 15(4), 453-464.
- Betz, N.E., & Weiss, D.J. (1973). An empirical study of computer administered two-stage ability testing (Research Report 73-4). Minneapolis, MN: University of Minnesota, Department of Psychology, Psychometric Methods Program.
- Betz, N.E., & Weiss, D.J. (1974). Simulation studies of two-stage ability testing (Research Report 74-4). Minneapolis, MN: University of Minnesota, Department of Psychology, Psychometric Methods Program.
- Bradshaw, G.D., & Gaudry, E. (1968). The effects of a single experience of success or failure on test anxiety. Australian Journal of Psychology, 20(3), 219-223.
- Brown, S.M., & Walberg, H.J. (1993). Motivational effects on test scores of elementary students. Journal of Educational Research, 86(3), 133-136.
- Frick, T.W. (1989). Bayesian adaptation during computer-based tests and computer-guided practice exercises. Journal of Educational Computing Research, 5(1), 89-114.
- Frick, T.W. (1990). A comparison of three decision models for adapting the length of computer-based mastery tests. Journal of Educational Computing Research, 6(4), 479-513.
- Gershon, R.C. & Bergstrom, B.A. (1991, April). Individual differences in computer adaptive testing: Anxiety, computer literacy and satisfaction. Paper presented at the annual meeting of the National Council of Measurement in Education, Chicago, IL.
- Hill, K.T. (1972). Anxiety in the evaluative context. In W.W. Hartup (Ed.), The young child (vol. 2) (pp. 225-263). Washington, DC: National Association for the Education of Young Children.
- Hill, K.T. (1980). Motivation, evaluation, and educational testing policy. In L.J. Fyans, Jr. (Ed.), Achievement motivation: Recent trends in theory and research. New York: Plenum Press.
- Kim, J.-G. (1993, November). Individual differences in computerized adaptive testing. Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans, LA.

- Kim, J.-G. (1995). The effects of computer adaptive testing and examinee test motivation on algebra skills, test anxiety, and attitudes. Unpublished doctoral dissertation, The University of Alabama, Tuscaloosa, AL.
- Kim, J.-G., McLean, J.E., & Moon, S.-B. (1994, November). A cross-cultural validation study of the computer attitude scale. Paper presented at the annual meeting of the Mid-South Educational Research Association, Nashville, TN.
- Kingsbury, G.G. (1990). Adapting adaptive testing: Using the MicroCAT testing system in a local school district. Educational Measurement: Issues and Practices, 9(2), 3-6, 29.
- Lapointe, A.E., Mead, N.A., & Phillips, G.W. (1989, January). A world of differences: An international assessment of mathematics and science (Report No. 19-CAEP.01). Princeton, NJ: Educational Testing Service.
- Legg, S.M., & Buhr, D.C. (1992). Computerized adaptive testing with different groups. Educational Measurement: Issues and Practice, 11(2), 23-27.
- Linn, R.L. (1993). Educational assessment: Expanded expectations and challenges. Educational Evaluation and Policy Analysis, 15(1), 1-16.
- Loyd, B.H., & Gressard, C.P. (1984). Reliability and factorial validity of computer attitude scales. Educational and Psychological Measurement, 44(2), 501-505.
- Martinez, M.E., & Mead, N.A. (1988). Computer competence: The first national assessment. Princeton, NJ: Educational Testing Service.
- McBride, J.R., & Martin, J.T. (1983). Reliability and validity of adaptive ability tests in a military setting. In D.J. Weiss (Ed.), New horizons in testing: Latent trait theory and computerized adaptive testing (pp. 223-236). New York: Academic Press.
- McKinley, R.L., & Reckase, M.D. (1980). Computer applications to ability testing. AEDS Journal, 13(3), 193-203.
- McMillan, J.H. (1980). Attitudes development and measurement. In J.H. McMillan (Ed.), The social psychology of school learning. New York: Academic Press.
- Mevarech, Z.R., & Ben-Artzi, S. (1987). Effects of CAI with fixed and adaptive feedback on children's mathematics anxiety and achievement. Journal of Experimental Education, 56(1), 42-46.
- Moe, K.C., & Johnson, M.F. (1988). Participants' reactions to computerized testing. Journal of Educational Computing Research, 4(1), 79-86.
- Moon, S.-B., Byun, C.-J., McLean, J.E., & Kaufman, A.S. (1994). Sequential-simultaneous profile analysis of Korean children's performance on the Kaufman assessment battery for children (K-ABC). Research in the Schools, 1(2), 29-35.

- Olsen, J.B., Maynes, D.D., Slawson, D., & Ho, K. (1989). Comparisons of paper-administered, computer-administered and computerized adaptive achievement tests. Journal of Educational Computing Research, 5(3), 311-326.
- Phillips, B.N., Pitcher, G.D., Worsham, M.E., & Miller, S.C. (1980). Test anxiety and the school environment. In I.G. Sarason (Ed.), Test anxiety: Theory, research, and applications (pp. 327-346). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pintrich, P.R. (1988). A process-oriented view of student motivation and cognition. In J.S. Stark & R. Mets (Eds.), Improving teaching and learning through research (pp. 55-70). San Francisco, CA: Jossey-Bass Publishers.
- Pintrich, P.R. (1989). The dynamic interplay of student motivation and cognition in the college classroom. In M. Maehr & C. Ames (Eds.), Advances in motivation and achievement: Motivation enhancing environments (Vol. 6, pp. 117-160). Greenwich, CT: JAI Press.
- Powell, Z. E. (1991). Test anxiety and test performance under computerized adaptive testing methods. Unpublished doctoral dissertation, Indiana University, Bloomington, IN.
- Rocklin, T., & O'Donnell, A.M. (1987). Self-adaptive testing: A performance-improving variant of computerized adaptive testing. Journal of Educational Psychology, 79(3), 315-319.
- Sarason, S.B., Davidson, K.S., Lighthall, F.F., Waite, R.R., & Ruebush, B.K. (1960). Anxiety in elementary school children: A report of research. New York: Wiley.
- Schmidt, F.L., Urry, V.W., & Gugel, J.F. (1978). Computer assisted tailored testing: Examinee reactions and evaluations. Educational and Psychological Measurement, 38(2), 265-273.
- Spielberger, C.D. (1980). Preliminary professional manual for the Test Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press.
- Urry, V.W. (1977). Tailored testing: A successful application of latent trait theory. Journal of Educational Measurement, 14(2), 181-196.
- Vispoel, W.P., & Coffman, D.D. (1992). Computerized adaptive testing of music-related skills. Bulletin of the Council for Research in Music Education, 112, 29-49.
- Vispoel, W.P., & Rocklin, T.R. (1993, April). Individual differences and test administration procedures: A comparison of fixed-item, adaptive, and self-adapted testing. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.
- Wainer, H. (1993). Measurement problems. Journal of Educational Measurement, 30(1), 1-21.

Weiss, D.J. (1982). Improving measurement quality and efficiency with adaptive testing. Applied Psychological Measurement, 6, 473-492.

Wise, S.L., Plake, B.S., Johnson, P.L., & Roos, L.L. (1992). A comparison of self-adapted and computerized adaptive tests. Journal of Educational Measurement, 29(4), 329-339.

• Wolf, L.F. (1993). The effects of motivation and anxiety on test performance. Dissertation Abstracts International, 54(4), 1334-A.

Wolf, L.F., Smith, J.K., & Birnbaum, M.E. (1994, April). Consequence of performance, test motivation, and mentally taxing items. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Yoes, M.E. (1994). TESTINFO: A computer program for computing test information using item response theory. St. Paul, MN: Assessment Systems Corporation.